

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-328586

(43)Date of publication of application : 13.12.1996

(51)Int.Cl.

G10L 3/02
G10L 3/00
H04N 5/928

(21)Application number : 07-130072

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 29.05.1995

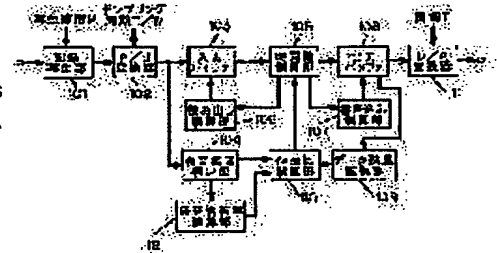
(72)Inventor : MISAHI MASAYUKI
NORIMATSU TAKESHI

(54) PHONETIC TIME AXIS CONVERSION DEVICE

(57)Abstract:

PURPOSE: To provide a phonetic time axis conversion device capable of converting an acoustic signal into a signal having the speech speed more than a reproducing speed and less than a recording time while holding an interval at the recording time when an acoustic signal is reproduced at the speed or below at the recording time.

CONSTITUTION: A sound M-multiple speed ($0 < M < 1$) reproduced from a recording/reproducing part 101 is written in an input buffer 103 after converted into a digital signal by an A/D converter at a period T/M . This signal is sound/silence-decided by a sound/silence decision part 104. A time axis control part 105 revises an elongation ratio according to the decision result of sound/silence parts to perform time axis elongation for the data read out from the input buffer 103 and to write in an output buffer. A remaining data monitor part 109 measures the remaining data stored without being read out from the output buffer to a D/A converter to impart them to an elongation ratio control part 110. The elongation ratio control part decides independently the elongation ratios of a sound part to a silence part by a conversion rule answering to the remaining data and silence contents to impart to the time axis control part.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-328586

(43) 公開日 平成8年(1996)12月13日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/02			G 1 0 L 3/02	A
	3/00			B
H 0 4 N 5/928			H 0 4 N 5/92	J

審査請求 未請求 請求項の数 5 O L (全 8 頁)

(21) 出願番号 特願平7-130072

(22) 出願日 平成7年(1995)5月29日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 三崎 正之

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

(72) 発明者 則松 武志

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

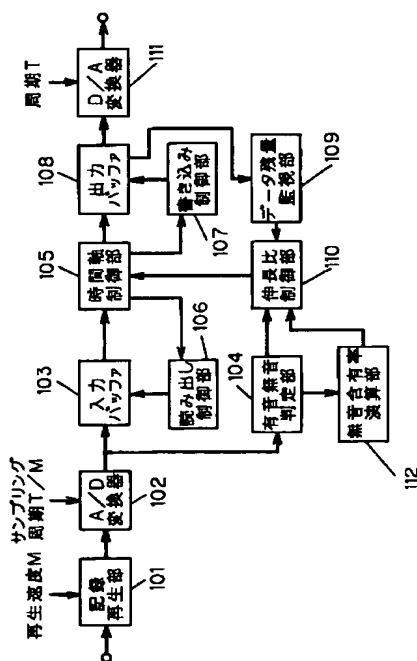
(74) 代理人 弁理士 滝本 智之 (外1名)

(54) 【発明の名称】 音声時間軸変換装置

(57) 【要約】

【目的】 音響信号を記録時の速度以下で再生する際、記録時の音程を保ちつつ再生速度以上で記録時以下の話速に変換する音声時間軸変換装置を提供する。

【構成】 記録再生部101からM倍速再生 ($0 < M < 1$) された音声は、A/D変換器102で周期T/Mでデジタル信号に変換されたのちに入力バッファ103に書き込まれる。この信号は有音無音判定部104で有音無音判定される。時間軸制御部105では入力バッファ103から読み出したデータに対して、有音無音部分の判定結果に応じて伸長比を変更して時間軸伸長を行い、出力バッファ108に書き込まれる。データ残量監視部109は、出力バッファからD/A変換器に読み出されずに蓄積されているデータの残量を計測し、伸長比制御部110に与える。伸長比制御部は、データ残量と無音含有率に対応した変換規則で有音部分と無音部分の伸長比を独立に決定して時間軸制御部に与える。



【特許請求の範囲】

【請求項 1】記録媒体に記憶された音響信号を記録時の M 倍（ただし、 $0 < M < 1$ ）の速度で読み出す再生部と、

前記再生部で読み出されたアナログ信号をデジタル信号へ変換する A/D 変換器と、

前記 A/D 変換器の出力データを記憶する入力バッファと、

前記入力バッファに入力される信号の有音部分と無音部分の判定を行う有音無音判定部と、

前記有音無音判定部の判定結果に基づき所定の時間長における無音部分の含有率を演算する無音含有率演算部と、

前記入力バッファに記憶されている音声データに時間軸伸長を行う時間軸制御部と、

前記時間軸制御部の出力データを蓄積するための出力バッファと、

前記出力バッファ内に蓄積されているデータのデータ残量を計測するデータ残量監視部と、

前記データ残量監視部から得られたデータ残量と前記無音含有率演算部から得られた無音含有率に応じて、予め

定めた規則にしたがって時間軸変換の伸長比を決定する伸長比制御部と、

前記出力バッファに蓄積されている音声データをアナログ信号に変換する D/A 変換器とを備えたことを特徴とする音声時間軸変換装置。

【請求項 2】伸長比制御部は、データ残量監視部から得られたデータ残量による短期的な適応制御と、無音含有率演算部から得られた無音含有率による長期的な適応制御、の双方を考慮した圧伸比の制御を行い、無音部分に

対する伸長比と有音部分に対する伸長比とを各々独立に設定することを特徴とする請求項 1 記載の音声時間軸変換装置。

【請求項 3】伸長比制御部は、無音部分伸長比を $1/M$ 以上に設定し、有音部分伸長比を 1.0 以上かつ $1/M$ 以下に設定し、データ残量と無音含有率に対応した変換規則に基づき各々の伸長比を決定することを特徴とする請求項 2 記載の音声時間軸変換装置。

【請求項 4】伸長比制御部は、有音部分伸長比を、データ残量が所定値以下の場合には $1/M$ に設定しそれ以外の場合には指定した固定値に設定し、無音部分伸長比を $1/M$ 以上の範囲でデータ残量と無音含有率に対応した変換規則に基づき伸長比を決定することを特徴とする請求項 2 記載の音声時間軸変換装置。

【請求項 5】伸長比制御部は、長期的な時間間隔で求めた無音含有率に応じて、有音または無音部分の伸長比の可変範囲の上限値または下限値を変更したデータ残量と無音含有率に対応した変換規則に調整して各々の伸長比を決定することを特徴とする請求項 3 記載の音声時間軸変換装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、ビデオテープレコーダー（VTR）等で音声の低速再生を行なう際に必要となる、音声の時間軸の長さを任意に伸長を行なうことを可能にする音声時間軸変換装置に関するものである。

【0002】

【従来の技術】所定の速度で記録された音声信号を、記録時の速度と異なる速度で再生する音声時間軸変換装置が実用化されている。例えば、テープレコーダーではテープの走行スピードを調節して再生速度を早めたり遅めたりすることができる。しかし、再生スピードが変化するのに伴って同時に音程も変化してしまうので、内容が聞きづらくなってしまう。そこで、音程を変化させずに再生速度のみを変化させることが可能な音声時間軸変換装置が提案されている。

【0003】以下、このような従来の音声時間軸変換装置について図面を参照しながら説明する。

【0004】図 5 は従来の音声時間軸変換装置の構成を表すブロック図である。図 6 は従来の時間軸伸長処理を表す模式図である。図 5 において、1 は音響信号の記録および再生を行なう記録再生部、2 は再生されたアナログ信号をデジタル信号に変換する A/D 変換器、3 はデジタルデータを蓄積するためのバッファメモリ、4 はバッファメモリ 3 から読み出されたデジタル信号をアナログ信号に変換する D/A 変換器、5 はバッファメモリ 3 へのデータの書き込みを制御する書き込み制御部、6 はバッファメモリ 3 のデータの読み出しを制御する読み出し制御部である。

【0005】以上のように構成された音声時間軸変換装置について、以下にその動作を説明する。ここでは、記録媒体への記録速度以下で音声信号を再生する場合に、音程を記録時の状態に戻して再生する音声時間軸変換装置について説明する。

【0006】まず、記録再生部 1 において記録時の M 倍（ $0 < M < 1$ ）の速度で音響信号が再生される。ここで記録再生部とは、例えば VTR、テープレコーダー等の磁気テープに記録する装置の記録再生部を意図している。次に、記録再生部 1 から再生された音響信号は、再生速度に反比例したサンプリング周期 T/M で A/D 変換器 2 によりデジタル信号に変換される。ここで、T は記録時の音響信号について標準化定理を満足するサンプリング周期であり、M 倍速再生された音響信号の場合には、その $1/M$ の周期にする必要がある。

【0007】A/D 変換されたこれらのデジタル信号は、書き込み制御部 5 によって周期 T/M でバッファメモリ 3 に順次蓄積されていく。ここで、バッファメモリ 3 に蓄積された各デジタル信号を周期 T で読み出し再生すれば記録時の音程に復元できるが、出力信号を連続して出し続けるには入力信号データが不足し、時間的に空

白となる部分ができる。そのため、読み出し制御部6ではバッファメモリ3に蓄えられたデジタル信号を数10 msecのフレーム単位で2度繰り返して読み出しを行う部分を設けるようにして、不足するデータを補うようにする。そして、読み出し制御部6により読み出されたデジタル信号を、D/A変換器4によりサンプリング周期Tでアナログ信号に変換する。これら一連の処理により、音程を変化させずに音声時間軸変換が実現できる。ここで説明した、音程一定で速度のみを変換する技術については、例えば「会話の時間軸を圧縮／伸長するテープ・レコーダ」；小坂、横堀、藤田；日経エレクトロニクス（1976. 7. 26発行）に詳しく解説されている。

【0008】図6は上述した音声時間軸変換装置の1/2倍速の場合の処理例を示している。(a)は記録時のデータを示しており、(b)はバッファメモリに蓄えられていくデータの時間的位置を示している。(b)の各ブロックを2回づつ繰り返しながらサンプリング周期Tで再生したものが(c)のデータ列となり、これは(a)のデータ列と音程が同じであり時間軸が2倍のスケールになっている。

【0009】

【発明が解決しようとする課題】上述した従来例では、記録時の音程を保ち、音声速度は記録媒体の再生速度と同一であり、記録時より遅く変換されている。ここでVTR等で画像情報を詳細にゆっくりと見たい場合などにおいて、記録媒体の再生速度を遅くしていくと、従来の時間軸変換装置を用いた場合には音声速度も画像と同様に遅くなっていく。ここで、人間の会話速度を違和感無く可変できる幅については、0.75～1.5倍程度と言われている。したがって、主として画像情報を詳細にゆっくりと見たい場合に記録媒体の再生速度をあまり遅くすると、再生される音声速度が必要以上に遅くなりすぎて違和感が起こり、かえって聴き取りにくくなる。また、これを回避するために、現在の記録媒体の再生速度よりも早い速度で音声を取扱しようすると、再生すべき音声信号データが時間的に不足してしまう。この場合、不足する音声データ部分が定期的に発生し、この部分に無音データを挿入するなどの手段を用いたとしても、不連続な音声信号となり極めて不自然な再生音となってしまう。

【0010】本発明は上記課題を解決するもので、記録媒体から記録速度以下の再生速度で読み出した場合に、音声の速度は必要以上に遅くせず、不連続点を生じることなく、聴き取りやすい音声を聴取することが可能な音声時間軸変換装置を提供することを目的とする。

【0011】

【課題を解決するための手段】上記課題を解決するために、請求項1記載の音声時間軸変換装置は、記録媒体に記憶された音響信号を記録時のM倍（ただし、 $0 < M < 1$ ）の速度で読み出す再生部と、前記再生部で読み出さ

れたアナログ信号をデジタル信号へ変換するA/D変換器と、前記A/D変換器の出力データを記憶する入力バッファと、前記入力バッファに入力される信号の有音部分と無音部分の判定を行う有音無音判定部と、前記有音無音判定部の判定結果に基づき所定の時間長における無音部分の含有率を演算する無音含有率演算部と、前記入力バッファに記憶されている音声データに時間軸伸長を行う時間軸制御部と、前記時間軸制御部の出力データを蓄積するための出力バッファと、前記出力バッファ内に蓄積されているデータのデータ残量を計測するデータ残量監視部と、前記データ残量監視部から得られたデータ残量と前記無音含有率演算部から得られた無音含有率に応じて、予め定めた規則にしたがって時間軸変換の伸長比を決定する伸長比制御部と、前記出力バッファに蓄積されている音声データをアナログ信号に変換するD/A変換器とを備えたものである。

【0012】請求項2記載の音声時間軸変換装置は、請求項1記載の構成において、無音部分に対する伸長比と有音部分に対する伸長比とをデータ残量による短期的な適応制御と、無音含有率による長期的な適応制御の双方を考慮して各々独立に設定する伸長比設定部を備えたものである。

【0013】請求項3記載の音声時間軸変換装置は、請求項2記載の発明において、無音部分伸長比を $1/M$ 以上に設定し、有音部分伸長比を 1.0 以上かつ $1/M$ 以下に設定し、データ残量に対応した変換規則に基づき各々の伸長比を決定する伸長比制御部を備えたものである。

【0014】請求項4記載の音声時間軸変換装置は、請求項2記載の発明において、有音部分伸長比をデータ残量が所定の値以下の場合には $1/M$ に設定し、それ以外には指定した固定値に設定し、無音部分伸長比を $1/M$ 以上の範囲でデータ残量に対応した変換規則に基づき伸長比を決定する伸長比制御部を備えたものである。

【0015】請求項5記載の音声時間軸変換装置は、請求項3記載の発明において、伸長比制御部は、長期的な時間間隔で求めた無音含有率に応じて、有音または無音部分の伸長比の可変範囲の上限値または下限値を変更したデータ残量と無音含有率に対応した変換規則に調整して各々の伸長比を決定する伸長比制御部を備えたものである。

【0016】

【作用】上記の構成によれば、有音無音の判定結果をもとに、無音部分の伸長比を有音部分より大きくした時間軸圧縮を行った後に出力バッファに書き込みを行う。この際に、出力バッファ内に蓄積されているデータのデータ残量を計測し、データ残量が少なくなるほど伸長比を大きくし、あるいは、無音部分の割合が所定値より少ない場合にも自動的に伸長比を加減してバッファメモリに絶えず十分なデータが確保される構成にしたことによ

り、時間的に短期的な伸長比の適応制御を行う。一方、無音含有率を比較的長い周期で更新することにより、入力される音声信号の性質が考慮され、時間的に長期的な伸長比の適応制御を行うことができる。これらの双方の伸長比の適応制御を行うことにより、有音部分の再生速度をできるだけ記録時に近い値に保って再生することができるとともに、入力信号の性質に合わせた滑らかな伸長比の変化を行うことができ、その結果、聴き取りやすいスロー再生音を得ることができる。

【0017】また、請求項4記載の構成によれば、バッファメモリに残っているデータの数であるデータ残量が極めて少ない場合には有音部分でも1/Mの伸長比で時間軸伸長して音切れを防ぎ、それ以外の場合には無音部分の伸長比をデータ残量をもとに調整する。これによって、話速は所定の固定値で再生しつつ、バッファメモリが空になることによって出力信号がとぎれることもない違和感の無い自然な再生音を得ることができる。

【0018】

【実施例】以下、本発明の第1の実施例について図面を参照しながら説明する。

【0019】図1は本発明の第1の実施例における音声時間軸変換装置の構成を示すブロック図である。図1において、101は音響信号の記録および再生を行なう記録再生部、102は記録再生部1で再生されたアナログ信号をデジタル信号に変換するA/D変換器、103はA/D変換された音響信号を一旦記録するための入力バッファ、104は入力バッファ103へ書き込まれるデジタル信号列が有音部分であるか無音部分であるかを判定する有音無音判定部、105は、入力バッファ103から読み出された信号に対して所定の伸長比で時間軸伸長処理を行う時間軸制御部、106は入力バッファ103からのデータの読み出しおよびそのアドレスを制御する読み出し制御部、107は出力バッファ108へのデータの書き込みおよびそのアドレスを制御する書き込み制御部、108は時間軸制御部105で処理されたデータを一時的に蓄える出力バッファ、109は出力バッファ108に一時的に保存しているデータ量を監視するデータ残量監視部、110は時間軸制御部105の伸長比をデータ残量監視部109の出力および無音含有率演算部112の出力に応じて決定する伸長比制御部、111は出力バッファ108に蓄積されたデジタルデータをアナログ信号に変換するD/A変換器、112は入力信号の無音含有率を演算する無音含有率演算部である。

【0020】以上のように構成された音声時間軸変換装置について、以下その動作を図1を参照しながら詳細に説明する。

【0021】まず、記録再生部101から記録時のM

($0 < M < 1$) 倍の速度で音響信号が読み出される。以後、速度とは記録速度に対する相対速度を表すこととする。ここで、記録再生部101よりM倍速で再生された

音響信号は、逐次A/D変換器102によりサンプリング周期T/Mでデジタル信号系列に変換されて、入力バッファ103に書き込まれる。一方、D/A変換器111はサンプリング周期Tでアナログ信号への変換が行われるので、単位時間あたり入力信号に比べて1/M倍以上の数のデータが出力バッファに適宜準備されていなければならない。その際に、入力信号全体に同じ割合の時間軸伸長を施すのではなく、有音無音判定部104による有音無音判定結果に基づき、無音部分には有音部分より大きい伸長比で時間軸伸縮することで有音部分の伸長比を1.0に近づける（有音部分のみ通常速度に近い速度で再生する）ように動作させるのが本実施例の基本的な考え方である。

【0022】入力バッファ103へ書き込まれる信号系列に対して、有音無音判定部104によりそのサンプル値列が有音部分であるか有音部分であるかの判定が行われる。この有音無音判定は、例えばサンプル値の音圧レベルが予め設定した基準音圧レベル以上であるか否かを判別することなどで実現できる。この判定結果は、入力バッファ103に書き込まれているデータのどの部分が有音/無音なのかを伸長比制御部110に伝えられると共に、無音含有率演算部112で、ある特定時間内における無音部分の含まれる割合を比較的長い時間周期で求めるために利用される。

【0023】時間軸制御部105は、後述する伸長比制御部によって与えられた伸長比で、入力バッファ103から読み出したデータに対して時間軸伸長処理を施して出力バッファ108へ出力する。その際には、無音部分には無音部分用の伸長比で時間軸伸長し、有音部分には有音部分用の伸長比で時間軸伸長が行われる。データ残量監視部109では、出力バッファ108に書き込まれているがD/A変換器111にはまだ出力されていないデータの残量をモニタしてその結果を伸長比制御部110に与えている。伸長比制御部110は、“無音含有率”と“データ残量”を入力とし、時間軸制御部105へ無音および有音の各々の部分に対する伸長比を与える。伸長比を与える基本的な考え方は以下の点である。

(1) 無音部分への伸長比を有音部分への伸長比より大きくまたは等しくする。

(2) 出力バッファへ溜まったデータ残量に応じて伸長比を調整する。その際、データ残量が少ないほど、伸長比を大きくする。

(3) 入力バッファへ書き込まれるデータの無音含有率に応じて伸長比を調整する。その際無音含有率が小さいほど、有音伸長比が小さめの値を取り易く、無音伸長比が大きめの値を取り易くする。

【0024】図2(a)には、データ残量と伸長比の関係、および無音含有率Rmによる曲線の変化させる傾向を示している。上記(1)は、一般的に人の音声は無音部分の継続時間長を調整して発話速度を変化させている

ことに基づいている。このことは、例えば、「『連続音声の中の音韻区分の持続時間の性質』；比企，金森，大泉；電気通信学会誌，第50巻，5号」に詳しく述べられている。また、(2)のように伸長比を変えることによって、出力バッファに残っているデータ残量が空に近くなるほど伸長比を大きくして出力データが不足しないようにし、データ残量が大きくなるほど伸長比を小さめに

してデータが溜まりにくくすることができる。この動作は、文節程度の時間長すなわち比較的短期的な適応の時定数にする。(3)は、例えば有音用伸長比は下に凸になる度合を大きくし、無音用伸長比は上に凸なる度合を大きくすることで実現される。

【0025】また、入力信号中に含まれる無音の量に応じて、有音部分と無音部分への伸長比を調整して、有音部分の伸長比をできるだけ1.0に近づけるようにする

目的で利用する。すなわち、無音の多く含まれるソースほど、有音部分への伸長比を1.0に近づけ易くして、相対的に有音部分を聴き取り易くする。この動作は、文章程度の時間長、すなわち比較的長期的な適応の時定数にする。そして、(2)および(3)を組み合わせることで、様々な音声信号に対して柔軟に適応することが可能となる。

【0026】データ残量と伸長比の関係は例えば図2

(a)のように関数形で与えられるものでも、あるいは階段状に変化するテーブル参照型のものでもかまわない。図2(b)の例は、有音部分はデータ残量が0にならない限り伸長率0、すなわち記録時と同一の話速で再生されることになる。この場合、有音部分の伸長率が固定の“1”の状態では、有音部分が連続すると出力バッファ内のデータ残量が急激に減少することになるので、

無音部分の伸長率はおおむね大きめにして出力バッファにデータが溜まりやすくしている。時間軸伸長することにより、出力バッファが空にならないようにデータ数を増加させることはできるが、むやみに大きい値の伸長比を与えていると出力バッファの容量を越えてしまうことになり、出力信号の連続性を保てなくなる。このため、データ残量が多くなるに連れて、無音部分の伸長比は小さく抑えてある。

【0027】以下は、記録媒体の再生速度を記録時の2/3倍(M=2/3)にした場合の動作について説明を

行う。

【0028】まず、図2の伸長比設定テーブルは、データ残量が0のとき、有音部分の伸長率を1.5にして、入力信号に有音が与えられても出力バッファが空になることを防いでいる。また、データ残量がほぼ出力バッファ容量と等しくなり、バッファ容量をオーバーフローする可能性がある場合には、無音部分の伸長比は1.5以下に抑える必要がある。

【0029】図4は、無音部分と有音部分とを別々の時間軸伸長比で時間軸伸長を行う場合の処理の様子を、時

間軸に関して模式的に示したものである。図4(a)の記録時の入力信号に対して図4(b)は2/3倍の再生速度で記録媒体から音声を再生した場合である。ここで、入力信号の無音部分の割合に依存して無音部分、有音部分の伸長比を決める必要がある。(c)と(d)には無音部分の割合の異なる2つの例を示す。入力信号1から6の部分において、(c)の例では1, 2, 3が無音部分で、4, 5, 6が有音部分とした場合の処理を行っている。(d)の例では1, 2が無音部分で、3, 4, 5, 6が有音部分とした場合の処理を行っている。この例では有音部分の伸長比はともに1.0にしているため、無音部分の伸長比は、(c)の例では2.0、(d)の例では2.5となる。

【0030】これらの例のように、無音部分の割合があらかじめ推定できれば、出力バッファ108から不足なく出力データをD/A変換器111に供給し続けられるので、伸長比を一定に固定しておいても出力データが不足することが無い。再生するソースの種類によって無音の含まれる割合は様々であるので、本実施例では、無音含有率を演算して、時間的に比較的長期な無音の含まれる割合を求めている。

【0031】しかしながら、時間的に比較的短期の間隔で考えると、出力バッファのデータ容量に制限がある場合や、遅延時間を考慮する必要がある場合などにおいては、局所的にデータが不足する場合が生じる。このような場合には、直接的に、出力バッファのデータ残量から短期的なデータ残量の増加率、すなわち伸長比を操作する必要がある。なぜなら、出力バッファから出ていくデータ量は一定値であるが、入力されるデータ量は有音部/無音部の伸長比によって増減率が変化するからであ

る。したがって、記憶容量に制限のある出力バッファに蓄えられたデータの量をモニタしその値によって伸長比を決定し、出力バッファで出力データの時間的な過不足を吸収することによって、無音の割合が予想できない音声であっても、無音部分と有音部分の伸長比を独立に設定してもデータが不足することがないようにすることができる。なお、本発明は時間軸処理の細かい部分に特徴があるわけでないため時間軸処理の詳細については言及しない。なお、時間軸変換処理の詳細については、例えば「『高品質音声速度変換方式のDSPによる実現』、鈴木，三崎；電子情報通信学会 音声研究会資料 SP90-34, (1990.8.23)」などに詳しく記述されている。

【0032】なお、図2の例では、有音部伸長比を1.0~1.5、無音部伸長比を1.5~3.0で変化させるとしたが、本発明においてこれらの伸長比の上限値・下限値はこれらの数値に限られるものではない。例えば、画像との時間的ずれを少なくするためには、伸長比の可変幅を小さくして、有音部伸長比を1.2~1.5、無音部伸長比を1.5~2.0などで変化させても良い。

【0033】また、図3に示したように、無音含有率が小さくなるに連れて、図3(a)から(b)、(c)へと順に変化させることも考えられる。これらの図は、可変する伸長比の上限値・下限値を無音含有率によって変更する例である。例えば、ニュースなどの比較的無音部分の多く含まれるソースの場合は、有音部分の伸長比の下限値を小さめの値に設定し、無音部分の伸長比の上限値も小さめの値に設定しても、データ残量が空になることはない。一方、サッカーなどの無音部分がほとんど含まれないソースでは、有音部分の伸長比の下限値を大きめにし、無音部分の伸長比の上限値を大きめにしても、データ残量が安定し易く、有音部分の安定した話速の再生音を得ることができる。このような場合においても、先に示した例と同様の効果が得られる。いずれの場合においても、平均的な有音部分の話速が安定し易くなる。

【0034】このような伸長比の制御を行なうことにより、無音部分の割合により時間軸伸長する伸長比が少々変化するが、記録時の話速以下で、かつ、記録媒体の再生速度より早い話速で、音声信号を聴取できることになる。

【0035】以上のように本実施例によれば、データ残量と無音含有率に基づいて有音部分・無音部分各々独立に時間軸伸長比を設定し、データ残量が予め定めた一定量より少ない時には有音部分の伸長比を $1/M$ に設定して出力信号が途切れることを防ぎつつ、有音部分をできるだけ記録時の話速に近くする伸長比の制御を行うことにより、記録媒体の再生速度が遅くなっても違和感なく聞き取りやすい再生音を得ることができる。

【0036】

【発明の効果】以上のように本発明は、記録速度の M 倍($0 < M < 1$)で再生された音響信号を有音無音判定部で有音部分と無音部分の判定をし、時間軸制御部では有音部分と無音部分に対して独立に設定した伸長比で時間軸伸長して出力バッファに蓄え、入力バッファに書き込まれる音声信号の無音含有率および出力バッファに蓄積されているデータ残量に応じて、無音部分および有音部分を時間軸伸長する伸長比を予め定めた規則により決定する伸長比制御部を備えているので、有音部分の再生速度をできるだけ記録時に近い値に保って再生することができるとともに、入力信号の性質に合わせた滑らかな伸長比の変化を行うことができ、その結果、聞き取りやすいスロー再生音を得ることができる。

【0037】そして、無音部分伸長比を $1/M$ 以上かつ、有音部分伸長比を 1.0 以上かつ $1/M$ 以下に設定して各々の伸長比を独立に変化させることにより、有音部分の話速をVTR再生速度より早くすることができる。

【0038】あるいは、有音部分伸長比はデータ残量が

所定の値以下の場合には $1/M$ に設定しそれ以外の場合には指定した固定値に設定し、かつ無音部分伸長比を $1/M$ 以上の範囲でデータ残量に対応した変換規則に基づいて決定することにより、話速は再生速度より早い一定値で再生することができる。したがって、一般的に有音部分の話速をより記録時に近い値にして出力できる。そして、無音含有率に応じて無音伸長比と有音伸長比の値、あるいは無音伸長比の値のみを調整できるように伸長比制御部で制御でき、入力信号の性質に合わせた伸長比に自動的に調整可能となる。また、出力バッファに溜まっているデータ残量を監視するデータ残量監視部を設けており、どのような入力信号が与えられても出力信号が途切れることなく再生できる。

【0039】また、長期的な時間間隔で求めた無音含有率に応じて、有音または無音部分の伸長比の可変範囲の上限値または下限値を変更したデータ残量と無音含有率に対応した変換規則に調整して各々の伸長比を決定することにより、入力信号の性質に適合した伸長比を設定することができる。

【0040】このように、画像信号をゆっくりと見るために記録媒体の再生速度を遅くした時に、どのような入力ソースであっても、音声信号を必要以上に遅い話速で聞く必要はなくなり、違和感の無い聞き取りやすいスロー再生を可能にする音声時間軸変換装置を提供することができる。

【図面の簡単な説明】

【図1】本発明の一実施例における音声時間軸変換装置の構成を示すブロック図

【図2】本実施例に係る伸長比設定テーブルの説明図

【図3】本実施例に係る他の伸長比設定テーブルの説明図

【図4】本実施例に係る時間軸伸長処理の模式図

【図5】従来の音声時間軸変換装置のブロック図

【図6】従来の時間軸伸長処理の模式図

【符号の説明】

- 101 記録再生部
- 102 A/D変換器
- 103 入力バッファ
- 104 有音無音判定部
- 105 時間軸制御部
- 106 読み出し制御部
- 107 書き込み制御部
- 108 出力バッファ
- 109 データ残量監視部
- 110 伸長比制御部
- 111 D/A変換器
- 112 無音含有率演算部

(a)

伸長比

3.0

2.0

1.5

1.0

0

N

データ残量

Rm:大

無音用伸長比

Rm:小

有音用伸長比

Rm:大

(b)

伸長比

3.0

2.0

1.5

1.0

0

N

データ残量Z

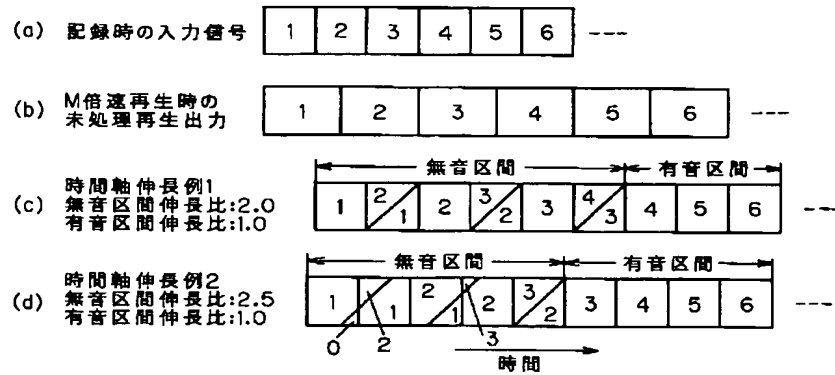
無音用伸長比

有音用伸長比

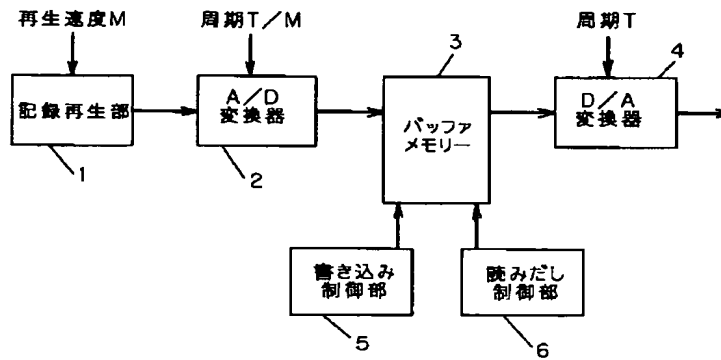
Figure 1 consists of three sub-graphs labeled (a), (b), and (c). Each graph plots the extension ratio (伸長比) on the vertical axis against the data residual amount (データ残量) on the horizontal axis. The horizontal axis for all graphs starts at 0 and ends at N.

- Graph (a):** The vertical axis ranges from 1.0 to 3.0. The '無音用伸長比' (extension ratio for non-speech data) is a solid line starting at 2.5 at 0 and decreasing to 1.5 at N. The '有音用伸長比' (extension ratio for speech data) is a solid line starting at 1.5 at 0 and decreasing to 1.0 at N. A horizontal dashed line is drawn at 1.5.
- Graph (b):** The vertical axis ranges from 1.2 to 3.0. The '無音用伸長比' (extension ratio for non-speech data) is a solid line starting at 3.0 at 0 and decreasing to 1.5 at N. The '有音用伸長比' (extension ratio for speech data) is a solid line starting at 1.5 at 0 and decreasing to 1.2 at N. A horizontal dashed line is drawn at 1.5.
- Graph (c):** The vertical axis ranges from 1.4 to 4.0. The '無音用伸長比' (extension ratio for non-speech data) is a solid line starting at 4.0 at 0 and decreasing to 1.5 at N. The '有音用伸長比' (extension ratio for speech data) is a solid line starting at 1.5 at 0 and decreasing to 1.4 at N. A horizontal dashed line is drawn at 1.5.

【図4】



【図5】



【図6】

